UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/814,081	03/31/2004	Wu Chou	633-001us	5517

47912   7590   11/18/2010
Avaya
DEMONT & BREYER, LLC
100 COMMONS WAY, STE 250
HOLMDEL, NJ 07733

EXAMINER
YEN, ERIC L

ART UNIT	PAPER NUMBER
2626	

NOTIFICATION DATE	DELIVERY MODE
11/18/2010	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

docketing@dblaw.com

	Application No.	Applicant(s)
Office Action Summary	10/814,081	CHOU ET AL.
	Examiner	Art Unit	
	ERIC YEN	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☒ Responsive to communication(s) filed on <u>20 September 2010</u>.
2a) ☐ This action is **FINAL.**   2b) ☒ This action is non-final.
3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) <u>1,3-7 and 9-18</u> is/are pending in the application.
    4a) Of the above claim(s) _____ is/are withdrawn from consideration.
5) ☐ Claim(s) _____ is/are allowed.
6) ☒ Claim(s) <u>1, 3-7, 9-18</u> is/are rejected.
7) ☐ Claim(s) _____ is/are objected to.
8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☐ The specification is objected to by the Examiner.
10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
    a) ☐ All   b) ☐ Some * c) ☐ None of:
      1. ☐ Certified copies of the priority documents have been received.
      2. ☐ Certified copies of the priority documents have been received in Application No. _____.
      3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08)
    Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Response to Amendment

1.   In response to the Final Office Action mailed 8/5/10, applicant has submitted an

amendment and Request for Continued Examination filed 9/20/10.

     Claims 1, 10, 12, 15, 18, have been amended.

Response to Arguments

2.   Applicant's arguments with respect to claims 1, 10, 12, 15, 18 have been

considered but are moot in view of the new ground(s) of rejection.

     Applicant has amended Claim 1 to recite "identifying, by the processor-based

device, the word received in the communication" and "utilizing by the processor-based

device a combination of terms that comprises:" (Amendment, page 13).

     However, as claimed the "identifying" and "utilizing" are not necessarily tied to

any particular purpose and therefore any form if "identification" and "use" for any reason

described in the prior art can read on the claim language.

     Applicant also argues that "the senses in Diab are not combined with words in

the same way as the distinctive combination" and "Diab teaches away from the

selection of terms according to Claim 1" because "nowhere does Diab disclose the

possibility of selecting both a word-term and a class-term [on whatever basis, not

necessarily IG] for classifying a word" And "Diab forecloses the possibility of selecting anything except a sense of a word" and "thus... Diab precludes the selection of a word-term because it only seeks to find a word class" (Amendment, pages 13-14).

Applicant argues that "the translated words of the target set are not combined with the candidate senses, because only the senses are candidates for selection, not the target words" and thus "Diab does not produce a combination like the distinctive combination of word-terms and word-classes recited by claim 1" and "Furthermore, Diab teaches away from the salient selection feature of the method of claim 1, because Daib seeks to select a sense, and only a sense, to be associated with a source word" and "in contrast, the method of claim 1 provides for the possibility of both a word-term and a word-class (or more than one) being selected relative to a word of interest" Amendment, pages 14-15).

First, Applicant's "teaching away" argument is also flawed because Diab only teaches something different done with word and word-class information. "Teaching away" applies where references teach away from other references, and some form of disparagement or criticism is required in order to constitute a "teaching away". Teaching of another use or alternative method is not sufficient to constitute a "teaching away".

Applicant also argues that Diab does not teach something that Diab was not applied to teach.

The relevant portion of the rejection states:

Diab teaches/suggests generating by the processor-based device a combination of terms, based on the word, wherein a term is one of a word-term and a word-class comprising: (i) a set of word-terms and (ii) a set of word-classes, wherein a term is one of a word-term and a word-class ("word sense tagging… automatically sense annotating… large amounts of data… using an unsupervised algorithm… bootstrap… creating a sense-tagged corpus", Introduction, especially paragraph 3; "project the sense tags from the target side to the source side… KIND-OF-DRAMA sense… CALAMITY… the tagging… would yield… large number of French words will receive tags from the English sense inventory", Approach, especially 4th bullet, paragraph ending at the upper right of page 257, and last paragraph; Applicant does not claim that the word that the generation is based on was derived from the communication, so as long as the generated data includes the word that also exists in the communication, it is "based on the word" that the communication "comprises"; Diab teaches sense-tagging words to create a corpus of sense-tagged data. Each of these sense-tagged words includes a word-class like CALAMITY and a word-term like catastrophe. Diab teaches the existence of different senses [like KIND-OF-DRAMA] and it is at least obvious that catastrophe is not the only word with different senses in French. The sense-annotated words in the corpus, collectively, are a "combination of terms" because the classes/senses are combined with their corresponding words, and at the very least in set theory sets can include 1 element, or alternatively, collectively all of the sense-annotated words are a set of word-terms and their corresponding senses collectively are a set of word-classes)

Nowhere in this portion of the rejection did Diab address the "selecting" that Applicant argues is not taught by Diab. <u>Lin</u> was applied to teach the "selecting" limitation. Diab was only applied to teach an <u>alternate source of word-term and word-class information</u>, particularly a sense-tagged corpus created by word tagging which includes words and word-classes. Created (i.e. "generated") sense-tagged corpora are "a combination of terms" "wherein a term is one of a word-term and a word-class comprising: (i) a set of word-terms and (ii) a set of word-classes" "wherein a term is one of a word-term and a word-class" because it is a single data entity including word-terms (i.e. words) and word-classes (senses, which are a category/class that provides semantic context for words). Applicant does not claim what a particular "combination" constitutes, and therefore any kind of combination that puts something that is a "word term" together with a "word class" in some way reads on the claim language (including the sense-tagged corpus data entity which was created to include words/word-terms and word senses/word-classes, and thus the words and word-classes were combined somehow to form the corpus data entity).

The rejection is based on a <u>combination</u> of references, and each reference contributes <u>a portion</u> of information to the combination. There is no requirement that one reference (i.e. Diab) teach each and every limitation, and therefore applicant's argument Diab does not teach the kind of "selecting" that applicant <u>intends</u> to claim is irrelevant because that was not what was Diab was applied to teach. Applicant did not address what the rejection stated Diab <u>did teach</u>.

The rejection asserts <u>Lin</u> performs the selection claimed, but Lin does not describe clearly where the terms/classes are <u>selected from</u> (we only know from Lin that they are selected from somewhere), and therefore Lin does not teach where the selection is from some <u>generated</u> data entity which includes word terms and word-classes.

Diab teaches a data entity/corpus which is <u>generated</u> and <u>includes</u> word-terms and word classes (words and word senses) which is the information that Lin selects.

The combination applied in the rejection therefore is where Lin performs selection of word-terms and word-classes from a created data entity that combines word terms and word-classes. Applicant did not address this combination and only attacked one reference in the combination. This is not sufficient to overcome a rejection.

Applicant then argues that a substitution would not be obvious because "Li-2002 does not disclose more than one corpus" and "since the present invention expressly discloses a combination of two different types of data to form the distinctive combination from which selections are made, there is no 'substitution' at issue" and "further, the Diab multi-step analysis would never word with a 'substitution' because it relies on mapping from one word to a target set to candidate senses" and "these concepts cannot reasonably be 'substituted' or combined with Li-2002" (Amendment, page 15).

As discussed above, both Li and Diab teach some data entity which includes the words and word-class information, only Li does not describe where the data entity is "generated" and Diab does teach where the data entity is "generated" (since a corpus is

created). Therefore, a substitution could be performed to obtain predictable results of selecting words and word-class information (as per Li) from a data entity that includes the words and word-class information (which Diab's sense-tagged corpus has). The prior art references <u>individually</u> do not need to teach <u>in detail</u> how selection from Diab's corpus is performed. Li explicitly teaches selection from something, and therefore one of ordinary skill in the art would recognize that some sort of analysis of a data entity to derive word-terms and word-classes is possible.

Therefore, the substitution is reasonable because the only difference between the <u>claimed</u> invention and Li is where the source of selection is a created data entity including the information selected. Diab teaches where the data entity including the information (which Li describes as selected) is known in the art, and one of ordinary skill in the art could substitute one <u>word-class/word-term source</u> with another.

At no point did the rejection state that Diab's "multi-step" analysis was substituted for any process in Li. Only the corpus, which is created and from which Li can select the terms and classes is used to substitute whatever source Li selects terms/classes from. Nor did the rejection state that Li-2002 disclosed more than one corpus, only that the terms and classes were selected from some <u>source of information including terms and classes</u>. It is this source of information which is substituted and Diab teaches where the <u>source</u> has the characteristics claimed (i.e. generated to include a combination of word terms and word classes).

Therefore, the examiner maintains similar prior art rejections to those previously presented.

Claim Rejections - 35 USC § 103

1.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

2.   Claims 1, 3, 5-7, 9-18, are rejected under 35 U.S.C. 103(a) as being

unpatentable over Li et al. ("Improving Latent Semantic indexing Based Classifier With

Information Gain"), hereafter Li, in view of Diab et al.("An Unsupervised Method for

Word Sense Tagging using Parallel Corpora"), hereafter Diab.

        As per Claim 1, 15, Li teaches a method (and corresponding apparatus, where

the joint classifier is defined in the claim by its function identical to the method in Claim

1) comprising: receiving, by a processor-based device, a communication that comprises

a word that is a natural-language word ("natural language understanding... directing the

user's call... matches a user's request", Section 4, Experimental Setup, paragraph 1;

where users communicate by speaking in natural language which includes speaking

natural language words)

        identifying, by the processor-based device, the word received in the

communication ("natural language understanding... directing the user's call... matches a

user's request", Section 4, Experimental Setup, paragraph 1; where to match a request

when a request is natural language speech, it is at least obvious that words

representing semantic information used to service the request are identified)

selecting by the processor-based device a plurality of terms wherein the selecting

is based on an information-gain value those terms that correspond to the word ("term-

document matrix… each selected term… IG based term selection is implemented…

terms are selected and used in the term-document matrix based on their discriminative

power", Section 3; where the terms selected are IG based and sorted by their individual

values. Applicant does not claim that <u>only</u> those terms that correspond to the word are

selected, and so selecting a larger set of terms which happens to include terms that

correspond to the word also reads on the claim language, and it is obvious that these

terms are selected because Li teaches categorizing the word is performed, which

cannot be done if the terms corresponding to the word were not selected)

generating by the processor-based device a matrix, wherein (i) the matrix

comprises a plurality of categories and a plurality of terms, and (ii) each term in the

matrix is associated with at least one category ("term-document matrix M is formed by

terms… selected term is mapped to a unique row vector and each category is mapped

to a unique column vector", Section 3, especially paragraph 3; where the matrix is

formed by combining terms [i.e. word terms] and categories [word-classes], and in a

matrix, the matrix cell corresponding to a specific row's term/word-term and a specific

column's category/word-class associates the term and the category corresponding to a

cell; Alternatively, words are naturally associated with some particular class [e.g. words

that are verbs, medical words, English words, etc.] and "at least one category" as

claimed does not necessarily refer to any category in particular so any word naturally

reads on this claim limitation because words are, by virtue of what they are, part of

some form of category)

determining from the matrix, by the processor-based device, a category for the

word ("LSI classifier... categorize an unknown document... derived from... as in LSI

according to IG enhanced term-document matrix... similarity... n-best categories",

Section 3; "user's request", Section 4; where categorizing a document by consequence

categorizes the word in that document, and this is "joint classification" in the sense that

it "jointly" uses both term information and category information to perform classification

[i.e. "joint classifier is configured to determine at least one category for the words, by

applying a combination of word information and word class information to the words",

Specification, page 6]).

Li fails to teach utilizing/generating by the processor-based device a combination

of terms that comprises: (i) a set of word-terms and (ii) a set of word-classes, wherein a

term is one of a word-term and a word-class, where the plurality of terms selected is

from the combination of terms, where those terms are in the combination of terms

Diab teaches/suggests generating by the processor-based device a combination

of terms, based on the word, wherein a term is one of a word-term and a word-class

comprising: (i) a set of word-terms and (ii) a set of word-classes, wherein a term is one

of a word-term and a word-class ("word sense tagging... automatically sense

annotating... large amounts of data... using an unsupervised algorithm... bootstrap...

creating a sense-tagged corpus", Introduction, especially paragraph 3; "project the

sense tags from the target side to the source side… KIND-OF-DRAMA sense…

CALAMITY… the tagging… would yield… large number of French words will receive

tags from the English sense inventory", Approach, especially 4th bullet, paragraph

ending at the upper right of page 257, and last paragraph; Applicant only claims

identification but does not specify where the identification limits anything else in the

claims, so as long as the generated data includes the word that also exists in the

communication, it is "based on the word" that the communication "comprises"; Diab

teaches sense-tagging words to create a corpus of sense-tagged data. Each of these

sense-tagged words includes a word-class like CALAMITY and a word-term like

catastrophe. Diab teaches the existence of different senses [like KIND-OF-DRAMA]

and it is at least obvious that catastrophe is not the only word with different senses in

French. The sense-annotated words in the corpus, collectively, are a "combination of

terms" because the classes/senses are combined with their corresponding words, and

at the very least in set theory sets can include 1 element, or alternatively, collectively all

of the sense-annotated words are a set of word-terms and their corresponding senses

collectively are a set of word-classes)

where the plurality of terms selected is from the combination of terms, where

those terms are in the combination of terms ("word sense tagging… automatically sense

annotating… large amounts of data… using an unsupervised algorithm… bootstrap…

creating a sense-tagged corpus", Introduction, especially paragraph 3; "project the

sense tags from the target side to the source side… KIND-OF-DRAMA sense…

CALAMITY… the tagging… would yield… large number of French words will receive

tags from the English sense inventory", Approach, especially 4th bullet, paragraph

ending at the upper right of page 257, and last paragraph; where Li [in Section 3,

especially paragraph 2] teaches that the term-document matrix is generated from a

labeled corpus, though it does not specifically state how the corpus is generated. Diab

teaches generating a labeled corpus by a processor which contains the very information

that Li wishes to extract for Li's matrix. Therefore, one of ordinary skill in the art can

simply substitute the corpus used in Li with one generated from a processor as per Diab

that contains the information needed to generate the matrix, and which can be used to

bootstrap the classifier described in Li)

Therefore, it would have been obvious to one of ordinary skill in the art at the

time of invention to perform a **simple substitution** of one corpus which the information-

gain-matrix performs selection from with another, because Li teaches a classification

method/device **which differed from the claimed device by the substitution of** a

corpus containing word class information and word term information generated by some

means used by Li to generated a matrix, with another corpus containing the same

information derived/generated by a processor. Diab teaches that a corpus generated by

a processor and containing word class information and word term information **was**

known in the art. One of ordinary skill in the art could have substituted one

corpus **for another to obtain the predictable results of** a system which performs

classification using a matrix generated from a corpus containing terms and classes (as

per Li) where the corpus containing terms and classes is generated by a processor (as

per Diab).

As per Claim 3, 16, 18, Li teaches/suggests (along with its apparatus equivalent of Claim 16 and article of manufacture equivalent in Claim 18, where claim 18 includes the limitations of both Claims 1 and 3 and so incorporates the rejections presented above regarding claim 1 as well) routing the communication by the processor-based device to a particular one of a plurality of destination terminals of a communication system, wherein the routing is based on the category of the word, and wherein the communication system comprises the processor-based device and the plurality of destination terminals ("routing… appropriate destination within a call center", Section 4; where the routing to destinations in the call-center/system based on the query's categorization, which is done via the joint classifier, which includes categorizing the words in the query, and the "system" can be interpreted as the call router and all of the places that the call is routed to; and this is "joint classification" in the sense that it "jointly" uses both term information and category information to perform classification [i.e. "joint classifier is configured to determine at least one category for the words, by applying a combination of word information and word class information to the words", Specification, page 6]).

As per Claim 5, Li teaches selecting of the plurality of terms is further based on a percentile value applied to the respective information-gain values of each term in the combination of terms ("top p percentile... according to the IG score", Section 3; where the terms being in the combination of terms is addressed in the same manner as above

by Diab in the parent claim, the set selected which is part of the corpus they are

selected from can be interpreted as the combination of terms as well).

As per Claim 6, Li teaches wherein the information-gain value for each term in

the combination of terms, indicates the average entropy variations over a plurality of

possible categories for each term in the combination of terms ("significance of the term

based on the entropy variations of the categories, which relates to the perplexity of the

classification task", Section 2; "literal terms… may not match those of a relevant

document", Section 1, paragraph 1; "IG enhanced… classified… categorize an unknown

document", Section 3; where the entropy variations are taught by Li to relate to

perplexity and so an entropy calculation is also a perplexity calculation and Equation 1

describes the information gain value being calculated from entropy/perplexity. Also the

subscript t_i at the end of Section 2 at least suggests that there is more than one term for

which the information gain is calculated; where the terms being in the combination of

terms is addressed above by Diab in the parent claim, the set selected which is part of

the corpus they are selected from can be interpreted as the combination of terms as

well).

As per Claim 7, 17, Li teaches (along with its apparatus equivalent of Claim 17)

wherein the category of the word is a cell in the term-category matrix ("cell… j-th

category", Section 3).

As per Claim 9, Li fails to teach wherein the combination of terms is generated by interleaving individual word-terms with their corresponding word-classes.

Diab teaches/suggests wherein the combination of terms is generated by interleaving individual word-terms with their corresponding word-classes ("word sense tagging... automatically sense annotating... large amounts of data... using an unsupervised algorithm... bootstrap... creating a sense-tagged corpus", Introduction, especially paragraph 3; "project the sense tags from the target side to the source side... KIND-OF-DRAMA sense... CALAMITY... the tagging... would yield... large number of French words will receive tags from the English sense inventory", Approach, especially 4th bullet, paragraph ending at the upper right of page 257, and last paragraph; where word-sense tagging interleaves [mixes or inserts the sense tags regularly between words in the corpus] the sense/word-classes and their respective words/word-terms)

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to perform a **simple substitution** of one corpus which the information-gain-matrix performs selection from with another, because Li teaches a classification method/device **which differed from the claimed device by the substitution of** a corpus containing word class information and word term information generated by some means used by Li to generated a matrix, with another corpus containing the same information derived/generated by a processor. Diab teaches that a corpus generated by a processor and containing word class information and word term information **was known in the art**. One of ordinary skill in the art could have substituted one corpus **for another to obtain the predictable results of** a system which performs

classification using a matrix generated from a corpus containing terms and classes (as

per Li) where the corpus containing terms and classes is generated by a processor (as

per Diab).

As per Claim 10, Li teaches/suggests a method comprising: receiving, by a

processor-based device, a communication that comprises at least one word, wherein

each of the at least one word that is a natural-language word ("natural language

understanding... directing the user's call... matches a user's request", Section 4,

Experimental Setup, paragraph 1; where users communicate by speaking in natural

language which includes speaking natural language words)

identifying, by the processor-based device, the word received in the

communication ("natural language understanding... directing the user's call... matches a

user's request", Section 4, Experimental Setup, paragraph 1; where to match a request

when a request is natural language speech, it is at least obvious that words

representing semantic information used to service the request are identified)

selecting by the processor-based device a plurality of terms wherein the selecting

is based on an information-gain value those terms that correspond to the word ("term-

document matrix... each selected term... IG based term selection is implemented...

terms are selected and used in the term-document matrix based on their discriminative

power", Section 3; where the terms selected are IG based and sorted by their individual

values)

generating by the processor-based device a term-category matrix, wherein (i) the term-category matrix comprises a plurality of terms and a plurality of categories, and (ii) each term in the matrix is associated with at least one category ("term-document matrix M is formed by terms... selected term is mapped to a unique row vector and each category is mapped to a unique column vector", Section 3, especially paragraph 3; where the matrix is formed by combining terms [i.e. word terms] and categories [word-classes], and in a matrix, the matrix cell corresponding to a specific row's term/word-term and a specific column's category/word-class associates the term and the category corresponding to a cell; Alternatively, words are naturally associated with some particular class [e.g. words that are verbs, medical words, English words, etc.] and "at least one category" as claimed does not necessarily refer to any category in particular so any word naturally reads on this claim limitation because words are, by virtue of what they are, part of some form of category)

classifying the communication by utilizing a joint classifier upon the at least one word, wherein the joint classifier comprises the term-category matrix ("LSI classifier... categorize an unknown document... derived from... as in LSI according to IG enhanced term-document matrix... similarity... n-best categories", Section 3; "user's request", Section 4; where categorizing a document by consequence categorizes the word in that document, and this is "joint classification" in the sense that it "jointly" uses both term information and category information to perform classification [i.e. "joint classifier is configured to determine at least one category for the words, by applying a combination of word information and word class information to the words", Specification, page 6]).

Li fails to teach generating by the processor-based device a combination of terms, based on the word, comprising: (i) a set of word-terms and (ii) a set of word-classes, wherein a term is one of a word-term and a word-class, where the plurality of terms selected is from the combination of terms, where those terms are in the combination of terms

Diab teaches/suggests generating by the processor-based device a combination of terms, based on the word, wherein a term is one of a word-term and a word-class comprising: (i) a set of word-terms and (ii) a set of word-classes, wherein a term is one of a word-term and a word-class ("word sense tagging… automatically sense annotating… large amounts of data… using an unsupervised algorithm… bootstrap… creating a sense-tagged corpus", Introduction, especially paragraph 3; "project the sense tags from the target side to the source side… KIND-OF-DRAMA sense… CALAMITY… the tagging… would yield… large number of French words will receive tags from the English sense inventory", Approach, especially 4th bullet, paragraph ending at the upper right of page 257, and last paragraph; Applicant does not claim that the word that the generation is based on was derived from the communication, so as long as the generated data includes the word that also exists in the communication, it is "based on the word" that the communication "comprises"; Diab teaches sense-tagging words to create a corpus of sense-tagged data. Each of these sense-tagged words includes a word-class like CALAMITY and a word-term like catastrophe. Diab teaches the existence of different senses [like KIND-OF-DRAMA] and it is at least obvious that catastrophe is not the only word with different senses in French. The sense-annotated

words in the corpus, collectively, are a "combination of terms" because the

classes/senses are combined with their corresponding words, and at the very least in

set theory sets can include 1 element, or alternatively, collectively all of the sense-

annotated words are a set of word-terms and their corresponding senses collectively

are a set of word-classes)

where the plurality of terms selected is from the combination of terms, where

those terms are in the combination of terms ("word sense tagging… automatically sense

annotating… large amounts of data… using an unsupervised algorithm… bootstrap…

creating a sense-tagged corpus", Introduction, especially paragraph 3; "project the

sense tags from the target side to the source side… KIND-OF-DRAMA sense…

CALAMITY… the tagging… would yield… large number of French words will receive

tags from the English sense inventory", Approach, especially 4th bullet, paragraph

ending at the upper right of page 257, and last paragraph; where Li [in Section 3,

especially paragraph 2] teaches that the term-document matrix is generated from a

labeled corpus, though it does not specifically state how the corpus is generated. Diab

teaches generating a labeled corpus by a processor which contains the very information

that Li wishes to extract for Li's matrix. Therefore, one of ordinary skill in the art can

simply substitute the corpus used in Li with one generated from a processor as per Diab

that contains the information needed to generate the matrix, and which can be used to

bootstrap the classifier described in Li)

Therefore, it would have been obvious to one of ordinary skill in the art at the

time of invention to perform a **simple substitution** of one corpus which the information-

gain-matrix performs selection from with another, because Li teaches a classification method/device **which differed from the claimed device by the substitution of** a corpus containing word class information and word term information generated by some means used by Li to generated a matrix, with another corpus containing the same information derived/generated by a processor. Diab teaches that a corpus generated by a processor and containing word class information and word term information **was known in the art**. **One of ordinary skill in the art could have substituted one** corpus **for another to obtain the predictable results of** a system which performs classification using a matrix generated from a corpus containing terms and classes (as per Li) where the corpus containing terms and classes is generated by a processor (as per Diab).

As per Claim 11, Li teaches wherein a cell I,j, of the term-category matrix represents a classification by the processor-based device of an i-th selected term into a j-th category ("LSI classifier... categorize an unknown document... similarity... n-best categories", Section 3; "user's request", Section 4; where categories in the matrix are among the j categories and categorizing a request including terms categorizes it into a category among the categories numbered by the values of j. Cells of matrices correspond to a particular row and common and in the case of a cell corresponding to a word/row and class/column, the cell represents an association between a word and the corresponding class)

As per Claim 12, Li teaches/suggests a method comprising: receiving, by a processor-based device, a communication that comprises a word that is a natural-language word ("natural language understanding... directing the user's call... matches a user's request", Section 4, Experimental Setup, paragraph 1; where users communicate by speaking in natural language which includes speaking natural language words)

identifying, by the processor-based device, the word received in the communication ("natural language understanding... directing the user's call... matches a user's request", Section 4, Experimental Setup, paragraph 1; where to match a request when a request is natural language speech, it is at least obvious that words representing semantic information used to service the request are identified)

selecting by the processor-based device a plurality of terms wherein the selecting is based on an information-gain value those terms that correspond to the word ("term-document matrix... each selected term... IG based term selection is implemented... terms are selected and used in the term-document matrix based on their discriminative power", Section 3; where the terms selected are IG based and sorted by their individual values)

wherein the selecting comprises: i) calculating an information gain value for each term that corresponds to the word ("terms are selected and used... according to IG criterion... sort the terms", Section 3; "terms in documents", Section 1, paragraph 1; where sorting terms by their IG values means that each term had its IG value calculated such that they can be sorted, and the terms are in documents that communicate

information [received at the input to the classification system], and documents contain words and so the terms in this context are words)

ii) sorting the terms in the union of terms in a descending order of information gain values ("sort the terms by their IG values in descending order", Section 3)

iii) setting a threshold of an information gain value corresponding to a specified percentile ("select top p percentile of terms according to the IG score distribution", Section 3; where taking the top p percentile sets the lowest of that p percentile as the threshold IG score)

iv) selecting only the terms having an information gain value greater than or equal to the threshold to generate the plurality of terms ("select top p percentile of terms", Section 3; where taking the top p percentile takes all terms exceeding the lowest IG value in that percentile and excludes everything falling below the percentile).

Li fails to teach generating by the processor-based device a combination of terms, based on the word, comprising: (i) a set of word-terms and (ii) a set of word-classes, wherein a term is one of a word-term and a word-class, where the plurality of terms selected is from the combination of terms, where the terms assigned information-gain values are the combination of terms, where the plurality of terms selected is from the combination of terms, where those terms are in the combination of terms, and where the terms are from the combination of terms

Diab teaches/suggests generating by the processor-based device a combination of terms, based on the word, wherein a term is one of a word-term and a word-class comprising: (i) a set of word-terms and (ii) a set of word-classes, wherein a term is one

of a word-term and a word-class ("word sense tagging… automatically sense

annotating… large amounts of data… using an unsupervised algorithm… bootstrap…

creating a sense-tagged corpus", Introduction, especially paragraph 3; "project the

sense tags from the target side to the source side… KIND-OF-DRAMA sense…

CALAMITY… the tagging… would yield… large number of French words will receive

tags from the English sense inventory", Approach, especially 4th bullet, paragraph

ending at the upper right of page 257, and last paragraph; Applicant does not claim that

the word that the generation is based on was derived from the communication, so as

long as the generated data includes the word that also exists in the communication, it is

"based on the word" that the communication "comprises"; Diab teaches sense-tagging

words to create a corpus of sense-tagged data. Each of these sense-tagged words

includes a word-class like CALAMITY and a word-term like catastrophe. Diab teaches

the existence of different senses [like KIND-OF-DRAMA] and it is at least obvious that

catastrophe is not the only word with different senses in French. The sense-annotated

words in the corpus, collectively, are a "combination of terms" because the

classes/senses are combined with their corresponding words, and at the very least in

set theory sets can include 1 element, or alternatively, collectively all of the sense-

annotated words are a set of word-terms and their corresponding senses collectively

are a set of word-classes)

where the plurality of terms selected is from the combination of terms, where

those terms are in the combination of terms, and where the terms are from the

combination of terms ("word sense tagging… automatically sense annotating… large

amounts of data… using an unsupervised algorithm… bootstrap… creating a sense-tagged corpus", Introduction, especially paragraph 3; "project the sense tags from the target side to the source side… KIND-OF-DRAMA sense… CALAMITY… the tagging… would yield… large number of French words will receive tags from the English sense inventory", Approach, especially 4th bullet, paragraph ending at the upper right of page 257, and last paragraph; where Li [in Section 3, especially paragraph 2] teaches that the term-document matrix is generated from a labeled corpus, though it does not specifically state how the corpus is generated. Diab teaches generating a labeled corpus by a processor which contains the very information that Li wishes to extract for Li's matrix. Therefore, one of ordinary skill in the art can simply substitute the corpus used in Li with one generated from a processor as per Diab that contains the information needed to generate the matrix, and which can be used to bootstrap the classifier described in Li)

Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to perform a **simple substitution** of one corpus which the information-gain-matrix performs selection from with another, because Li teaches a classification method/device **which differed from the claimed device by the substitution of** a corpus containing word class information and word term information generated by some means used by Li to generated a matrix, with another corpus containing the same information derived/generated by a processor. Diab teaches that a corpus generated by a processor and containing word class information and word term information **was known in the art**. **One of ordinary skill in the art could have substituted one**

corpus **for another to obtain the predictable results of** a system which performs

classification using a matrix generated from a corpus containing terms and classes (as

per Li) where the corpus containing terms and classes is generated by a processor (as

per Diab).

As per Claim 13, Li teaches wherein the selected terms in the plurality of terms

are processed by the processor-based device to form a term-category matrix from

which a joint classifier determines at least one category for the word, and wherein the

processor-based device comprises the joint classifier ("**LSI** classifier... categorize an

unknown document... derived from... as in **LSI** according to IG enhanced term-

document matrix... similarity... n-best categories", Section 3; "user's request", Section

4; where categorizing a document by consequence categorizes the word in that

document, and this is "joint classification" in the sense that it "jointly" uses both term

information and category information to perform classification [i.e. "joint classifier is

configured to determine at least one category for the words, by applying a combination

of word information and word class information to the words", Specification, page 6]).

As per Claim 14, Li teaches generating by the processor-based device a term-

category matrix, wherein (i) the term-category matrix comprises a plurality of terms and

a plurality of categories, and (ii) each term in the matrix is associated with at least one

category ("term-document matrix **M** is formed by terms... selected term is mapped to a

unique row vector and each category is mapped to a unique column vector", Section 3,

especially paragraph 3; where the matrix is formed by combining terms [i.e. word terms]

and categories [word-classes], and in a matrix, the matrix cell corresponding to a

specific row's term/word-term and a specific column's category/word-class associates

the term and the category corresponding to a cell; Alternatively, words are naturally

associated with some particular class [e.g. words that are verbs, medical words, English

words, etc.] and "at least one category" as claimed does not necessarily refer to any

category in particular so any word naturally reads on this claim limitation because words

are, by virtue of what they are, part of some form of category)

       selecting from the term-category matrix, by the processor-based device, a

category for the word ("LSI classifier... categorize an unknown document... derived

from... as in LSI according to IG enhanced term-document matrix... similarity... n-best

categories", Section 3; "user's request", Section 4; where categorizing a document by

consequence categorizes the word in that document, and this is "joint classification" in

the sense that it "jointly" uses both term information and category information to perform

classification [i.e. "joint classifier is configured to determine at least one category for the

words, by applying a combination of word information and word class information to the

words", Specification, page 6])

       routing the communication by the processor-based device to a particular one of a

plurality of destination terminals of a communication system, wherein the routing is

based on the category of the word, and wherein the communication system comprises

the processor-based device and the plurality of destination terminals ("routing...

appropriate destination within a call center", Section 4; where the routing to destinations

in the call-center/system based on the query's categorization, which is done via the joint

classifier, which includes categorizing the words in the query, and the "system" can be

interpreted as the call router and all of the places that the call is routed to).

1.   Claim 4 is rejected under 35 U.S.C. 103(a) as being unpatentable over Li, in view

of Diab, as applied to Claim 1, above, and further in view of Sakai et al. (US 7,099,819),

hereafter Sakai.

        As per Claim 4, Li, in view of Diab, fail to teach wherein an automatic word class

clustering algorithm is utilized to generate the word-classes.

        Sakai teaches wherein an automatic word class clustering algorithm is utilized to

generate the word-classes ("category decision rules... each text is classified to a

category according to the category decision rule", col. 3, lines 35-50; "automatically

creates a new category", col. 6, line 53 – col. 7, line 5; "if a cluster consisting of a large

number of texts... new category to which this cluster is classified", col. 6, lines 34-40;

"cluster generation unit", col. 6, lines 7-24; where the clustering is automatically

performed and whose results is used for a new word class, and so it is an automatic

word class clustering algorithm and is used to generate new word class [i.e., category]

rules/information. Li and Diab teach where categories/senses are generated somehow

[since they must have been derived from somewhere to be used in tagging], without

providing the specifics. Sakai teaches another method for generating the same data

and so a simple substitution of the generation can be performed to yield the class/sense information used in Diab's sense tagging).

Therefore, it would have been obvious to one of ordinary skill in the art to perform a **simple substitution** of one category with another, because Li and Diab teach **a device which differed from the claimed device by the substitution of** categories generated by some unspecified manner with categories generated by clustering. Sakai teaches that categories generated by clustering **were known in the art**. **One of ordinary skill in the art could have substituted one known element for another** by using categories generated from clustering instead in the unsupervised tagging in Diab **in order to obtain the predictable results of** a system that performs classification based on a matrix derived from word class and word term data (Li) where the word class and word term data are automatically generated by a processor (Diab) based on classes determined initially using a form of clustering (Sakai).

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to ERIC YEN whose telephone number is (571)272-4249. The examiner can normally be reached on M-F 7:30-4:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a

USPTO Customer Service Representative or access to the automated information

system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

EY 11/12/10
        /Eric Yen/
        Examiner, Art Unit 2626